

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2000-181843

(P2000-181843A)

(43)公開日 平成12年6月30日(2000.6.30)

(51)Int.Cl. ⁷	識別記号	F I	テーマコード(参考)
G 0 6 F 13/00	3 5 4	G 0 6 F 13/00	3 5 4 D 5 B 0 8 9
H 0 4 L 12/54		H 0 4 L 11/20	1 0 1 B 5 K 0 3 0
12/58			9 A 0 0 1

審査請求 未請求 請求項の数10 O L (全 19 頁)

(21)出願番号 特願平10-361124

(22)出願日 平成10年12月18日(1998.12.18)

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 松本 茂

東京都府中市東芝町1番地 株式会社東芝
府中工場内

(72)発明者 長谷川 義朗

東京都府中市東芝町1番地 株式会社東芝
府中工場内

(74)代理人 100058479

弁理士 鈴江 武彦 (外6名)

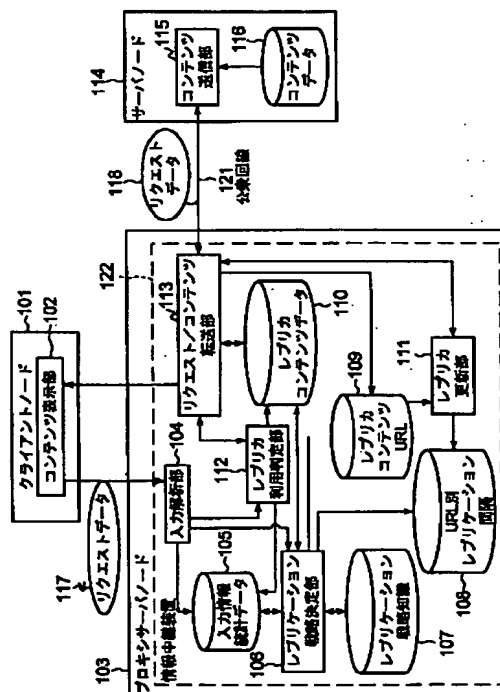
最終頁に続く

(54)【発明の名称】 情報中継装置及び記憶媒体

(57)【要約】

【課題】 本発明は、オリジナルコンテンツの更新頻度が高く、さらにクライアントから不特定多数のサーバへの情報提供要求が出力される場合において、その平均レスポンスタイムを向上させることができる。

【解決手段】 ネットワーク121上に配置され、クライアント計算機からのデータ提供要求をサーバ計算機へ転送し、サーバ計算機からの提供データをクライアント計算機へ転送する中継機能を有する情報中継装置122において、クライアント計算機から過去にデータ提供要求されたサーバ計算機上のデータについて、データ提供要求と非同期に、その要求データを複製し保持するレプリケーション手段113、110、109、111、108と、クライアント計算機からデータ提供要求されたとき、レプリケーション手段が保持する複製データ110をクライアント計算機に返す複製返却手段113とを備えた情報中継装置。



【特許請求の範囲】

【請求項 1】 情報提供者となるサーバ計算機と情報利用者となるクライアント計算機の間を繋ぐネットワーク上に配置され、クライアント計算機からのデータ提供要求をサーバ計算機へ転送し、サーバ計算機からの提供データをクライアント計算機へ転送する中継機能を有する情報中継装置において、

前記クライアント計算機から過去にデータ提供要求されたサーバ計算機上のデータについて、データ提供要求と非同期に、その要求データを複製し保持するレプリケーション手段と、

前記クライアント計算機からデータ提供要求されたとき、前記レプリケーション手段が保持する複製データを前記クライアント計算機に返す複製返却手段とを備えたことを特徴とする情報中継装置。

【請求項 2】 前記レプリケーション手段は、設定された時間間隔で複製データを最新状態に更新することを特徴とする請求項 1 記載の情報中継装置。

【請求項 3】 複数の複製データについての各々の更新時間間隔を、各データに対するクライアント計算機からのデータ提供要求の頻度に基づいて変更するレプリケーション間隔決定手段を備えたことを特徴とする請求項 2 記載の情報中継装置。

【請求項 4】 複数の複製データについての各々の更新時間間隔を、各データに関するサーバ計算機上でのデータ変更頻度に基づいて変更するレプリケーション間隔決定手段を備えたことを特徴とする請求項 2 又は 3 記載の情報中継装置。

【請求項 5】 前記データ提供要求を受けた場合に、前記レプリケーション手段が保持する複製データが有効期限内のものであるか否かを判断するレプリカ利用判定手段を備え、

前記レプリケーション手段は、前記レプリカ利用判定手段により複製データが有効期限外と判定された場合には、サーバ計算機に複製データに対応したデータの属性情報を問い合わせ、その属性情報から該当データがサーバ計算機上で更新されていると判断したときには、サーバ計算機からデータを取得し直すとともに、その取得データを前記中継機能に引き渡すことを特徴とする請求項 1 乃至 4 のうち何れか 1 項記載の情報中継装置。

【請求項 6】 前記サーバ計算機は、提供用のデータに変更が起きた場合、データ提供登録されたデータについてその登録を行ったデータ提供対象に、更新後のデータを送信するものである場合に、

前記レプリケーション手段は、クライアント計算機から複製データを保持していないデータについてのデータ提供要求を受けると、前記サーバ計算機に、要求されたデータをデータ提供するように登録することを特徴とする請求項 1 記載の情報中継装置。

【請求項 7】 前記レプリケーション手段は、前記複製

データとともにその複製データについてのアクセス制限情報を取得し保持するとともに、

前記複製返却手段は、前記アクセス制限情報に基づき、アクセス許可されるクライアントからのデータ提供要求にのみ複製データを返すことを特徴とする請求項 1 乃至 6 のうち何れか 1 項記載の情報中継装置。

【請求項 8】 情報提供者となるサーバ計算機と情報利用者となるクライアント計算機の間を繋ぐネットワーク上に配置され、クライアント計算機からのデータ提供要求をサーバ計算機へ転送し、サーバ計算機からの提供データをクライアント計算機へ転送する中継機能を有する情報中継装置を制御するプログラムであって、

コンピュータに、

前記クライアント計算機から過去にデータ提供要求されたサーバ計算機上のデータについて、データ提供要求と非同期に、その要求データを複製させ保持させるレプリケーション手段と、

前記クライアント計算機からデータ提供要求されたとき、前記レプリケーション手段が保持させる複製データを前記クライアント計算機に返させる複製返却手段とを有するプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【請求項 9】 前記レプリケーション手段は、設定された時間間隔で複製データを最新状態に更新させることを特徴とする請求項 8 記載の記憶媒体。

【請求項 10】 前記サーバ計算機は、提供用のデータに変更が起きた場合、データ提供登録されたデータについてその登録を行ったデータ提供対象に、更新後のデータを送信するものである場合に、

前記レプリケーション手段は、クライアント計算機から複製データを保持していないデータについてのデータ提供要求を受けると、前記サーバ計算機に、要求されたデータをデータ提供するように登録させることを特徴とする請求項 8 記載の記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は情報中継装置及び記憶媒体、更に詳しくは、ネットワーク上のサーバ計算機で生成される情報について、これを参照するクライアント計算機と間であって両者の情報授受を中継する情報中継装置及び記憶媒体に関する。

【0002】

【従来の技術】WWW (World Wide Web) やFTP (File Transfer Protocol) サービスに代表される分散計算機システムでは、一塊の情報 (以下、コンテンツと呼ぶ) を特定するユニークな識別文字列がURL という形でコンテンツ毎に関連付けられている。クライアントシステム (以下、クライアントという) からはURL を指定することでサーバシステム (以下、サーバという) からコンテンツを

引き出すことができる。

【0003】このようなシステムを利用する際、クライアントとサーバの間にプロキシ（Proxy）サーバと呼ばれる情報中継装置を介在させる場合がある。プロキシサーバはいくつかの機能を持つが、その1つにキャッシング機能がある。

【0004】このキャッシング機能を利用することにより、クライアントが一度参照したコンテンツがプロキシサーバを通過する際にそのキャッシュエリアに格納される。このデータキャッシュにより、クライアントから再びそのコンテンツを見たいという要求が出されたとき、キャッシュ内のデータを返すためプロキシサーバとサーバ間のネットワークデータ転送を省くことができ、クライアントの要求に高速に答えることができる。

【0005】現在実現されている一般的なキャッシングプロキシサーバの動作は以下ようになる。ここではWWWの例を示す。

【0006】1. クライアントからの表示要求を受け、キャッシュエリアを調べる。

【0007】2. キャッシュエリアにクライアントから要求されたコンテンツの複製が存在する場合には、複製が有効期限内であればそれを返す。有効期限は、サーバからコンテンツを取得する際に一緒に送られてくる。

【0008】3. 有効期限をすぎている場合、それが最新のものかどうかを確認するため、オリジナルが置かれているWWWサーバに当該コンテンツの最終更新日時を問い合わせる。

【0009】4. オリジナルの最終更新日時が複製の最終更新日時と一致した場合は、キャッシュ内のコンテンツをクライアントに返す。

【0010】5. オリジナルの最終更新日時がキャッシュより新しい場合は、改めてオリジナルのコンテンツをダウンロードし、クライアントに返すと共に、キャッシュを更新する。

【0011】以下、上記2または4の状況を「キャッシュがヒットした」と表現する。

【0012】ここで、クライアントからの情報提供要求に対し、キャッシュがヒットするための条件を整理すると以下ようになる。

【0013】・要求されたURLのコンテンツがキャッシュに存在する（以下、存在性という）。

【0014】・キャッシュに存在する当該コンテンツがサーバと同じものである（以下、最新性という）。

【0015】これらの条件の性質からして、キャッシュ上の複製データに存在性と最新性の両条件が成り立ったときのみ、キャッシュがヒットすることになる。

【0016】一方、データベースのレプリケーションシステムやプッシュ（Push）テクノロジーと呼ばれる技術では、あらかじめ設定したコンテンツに対し、サーバ上のコンテンツに内容更新が起こった時点で複製を持

つクライアントに変更が通知され、複製が最新状態に変更される。

【0017】プロキシサーバのキャッシュ管理にプッシュ技術を利用することにより、あらかじめ設定されたURLのコンテンツに対しては常に最新状態のものがキャッシュエリアに保持される。このため、一般的なプロキシサーバよりもクライアントからの要求に対するキャッシュのヒット確率が高くなり、クライアントから見た平均レスポンス時間が短くなる効果がある。サーバ上のコンテンツが頻繁に更新される場合、この効果が顕著になる。

【0018】また、厳密に言えばプッシュ技術とはいえないが、一般にプッシュ技術といわれているものにインテリジェントプル（pull）と呼ばれる技術がある。上記したプッシュ技術ではサーバ側から積極的にコンテンツを送るのに対し、このインテリジェントプルではクライアント側から定期的にコンテンツを取得することで見かけ上プッシュ技術を用いたのと同様な状態を実現しようとする。

【0019】このインテリジェントプル技術をプロキシサーバに適用した例として、米国Marimba社のCastanetがある。

【0020】Marimba社のCastanetはサーバ機能を持つトランスミッタとクライアント機能を持つチューナから構成される。チューナはあらかじめチューナ上で設定したURLのコンテンツがトランスミッタ上で更新されたかを定期的にチェックし、更新があったコンテンツをダウンロードして、チューナ上の複製ファイルを更新する。

【0021】ユーザからコンテンツの参照要求があると、チューナ上の複製コンテンツを表示する。したがって、チューナをWWWブラウザのプロキシサーバとして利用すると、高速なレスポンスが実現できる。この場合、WWWブラウザからの参照要求が、チューナ上にキャッシュされているコンテンツなら複製が返され、そうでない場合はオリジナルのサーバからコンテンツをダウンロードする。

【0022】チューナにキャッシュされているコンテンツは定期的に最新状態に更新されているため、一般的なプロキシサーバを利用する場合に比べ、キャッシュの最新性が成立する確率が高くなる。

【0023】また、特開平10-21134には、2つのプロキシのプロセスを用意しておき、1つ目のプロキシのプロセスは通常のプロキシの機能を有し、2つ目のプロキシは1つ目のプロキシが生成するアクセスログに記載されているURLを順にWebサーバにあるオリジナルから読み込んでキャッシュに記憶させる技術が記載されている。つまり、プロキシのキャッシュにクライアントが取得しようとしている目的のコンテンツが見つからない場合、もしくは見つかったも有効期限が切れてい

る場合はオリジナルを読み込んできて、目的のコンテンツで、有効期限内のコンテンツを見つけた場合は、そのままクライアントに返すのである。

【0024】

【発明が解決しようとする課題】ところが、上述のキャッシングプロキシサーバのキャッシング機能を用いた場合は、サーバに置かれているオリジナルコンテンツの内容更新頻度が低い場合、キャッシングプロキシサーバのキャッシュ機能は有効に働く。しかし、オリジナルコンテンツの更新頻度が高い場合、キャッシュの最新性が成立する確率は低くなるため、キャッシュのヒット率が下がり、プロキシサーバがキャッシングする効果が低くなるという問題がある。

【0025】また、上記プッシュ技術やインテリジェントプル技術を用いた場合においては、あらかじめ設定されたコンテンツしか複製が作られない。したがって、WWWのようにオリジナルコンテンツが不特定のシステムではキャッシュの存在性が成立する確率が下がり、ヒット率が低くなる問題がある。仮に、十分多くのコンテンツをあらかじめプッシュ対象に設定することによって存在性を向上させた場合、必要なキャッシュの大きさが大きくなりすぎるという問題が発生し、現実的とはいえない。

【0026】さらに、特開平10-21134記載の技術では、コンテンツをキャッシュするタイミングが、1つ目のプロキシのアクセスログが増加するタイミングでアクセスログに基づいて行われるため、クライアントからのリクエストが無い限り、キャッシュの更新は行われず、有効期限内にオリジナルのコンテンツが変更されていても、クライアントには新しいコンテンツをクライアントが入手できない場合が生じる。

【0027】本発明は、このような実情を考慮してなされたもので、オリジナルコンテンツの更新頻度が高く、さらにクライアントから不特定多数のサーバへの情報提供要求が出力される場合において、その平均レスポンスタイムを向上させることができる情報中継装置及び記憶媒体を提供することを目的とする。

【0028】

【課題を解決するための手段】上記課題を解決するために、請求項1に対応する発明は、情報提供者となるサーバ計算機と情報利用者となるクライアント計算機の間を繋ぐネットワーク上に配置され、クライアント計算機からのデータ提供要求をサーバ計算機へ転送し、サーバ計算機からの提供データをクライアント計算機へ転送する中継機能を有する情報中継装置についてなされたものである。

【0029】この情報中継装置においては、レプリケーション手段により、クライアント計算機から過去にデータ提供要求されたサーバ計算機上のデータについて、データ提供要求と非同期に、その要求データが複製され保

持される。

【0030】また、クライアント計算機からデータ提供要求されたときには、複製返却手段によって、レプリケーション手段が保持する複製データがクライアント計算機に返される。

【0031】本発明では、このようなに、データ提供要求とは非同期な複製更新を行うことから、オリジナルコンテンツの更新頻度が高い場合でも適切な複製返却を可能とすることができ、また、過去に要求されたデータについて、複製を行うようにしているので、クライアントから不特定多数のサーバへの情報提供要求が出力される場合にその平均レスポンスタイムを向上させることができる。

【0032】次に、請求項2に対応する発明は、請求項1に対応する発明において、レプリケーション手段は、設定された時間間隔で複製データを最新状態に更新する。

【0033】したがって、より効率的に、複製データの返却を行うことができる。

【0034】次に、請求項3に対応する発明は、請求項2に対応する発明において、レプリケーション間隔決定手段が設けられ、同間隔決定手段によって、複数の複製データについての各々の更新時間間隔が、各データに対するクライアント計算機からのデータ提供要求の頻度に基づいて変更される。

【0035】したがって、データ要求頻度に応じて適切な複製更新を行うことができる。

【0036】次に、請求項4に対応する発明は、請求項2又は3に対応する発明において、レプリケーション間隔決定手段によって、複数の複製データについての各々の更新時間間隔が、各データに関するサーバ計算機上でのデータ変更頻度に基づいて変更される。

【0037】したがって、クライアントの要求のみならず、サーバ側での状況をも考慮した複製更新を行うことができる。

【0038】次に、請求項5に対応する発明は、請求項1～4に対応する発明において、レプリカ利用判定手段によって、データ提供要求を受けた場合に、レプリケーション手段が保持する複製データが有効期限内のものであるか否かが判断されるようになっている。

【0039】このレプリカ利用判定手段により複製データが有効期限外と判定された場合には、さらに、レプリケーション手段により、サーバ計算機に複製データに対応したデータの属性情報が問い合わせられ、その属性情報から該当データがサーバ計算機上で更新されていると判断されたときに、サーバ計算機からデータが取得し直されるとともに、その取得データが前記中継機能に引き渡される。

【0040】したがって、実際にはサーバ側のデータが変更されていない場合に、サーバ～中継装置間で重複し

てデータ送信されることを防止することができ、効率的なデータ転送を行うことができる。

【0041】次に、請求項6に対応する発明は、請求項1に対応する発明において、サーバ計算機は、提供用のデータに変更が起きた場合、データ提供登録されたデータについてその登録を行ったデータ提供対象に、更新後のデータを送信するものである場合に情報中継装置に適用されるものである。なお、このサーバ計算機は例えばプッシュ技術が使用されたものである。

【0042】この場合に、レプリケーション手段は、クライアント計算機から複製データを保持していないデータについてのデータ提供要求を受けると、サーバ計算機に、要求されたデータをデータ提供するよう登録する。

【0043】したがって、情報中継装置からプッシュ技術等を用いたサーバ計算機にプッシュ対象データを登録するよう働きかけることで、請求項1の発明と同様な効果を奏することができる。

【0044】次に、請求項7に対応する発明は、請求項1に対応する発明において、レプリケーション手段によって、複製データとともにその複製データについてのアクセス制限情報が取得され保持されるようになってい

る。

【0045】さらに、複製返却手段によって、アクセス制限情報に基づき、アクセス許可されるクライアントからのデータ提供要求にのみ複製データが返されるようになっている。

【0046】したがって、アクセス権限を有するクライアントに対してのみ、データ返却することが可能となる。

【0047】次に、請求項8に対応する発明は、請求項1に対応する発明をコンピュータに実現させるプログラムを記録した記録媒体である。

【0048】この記録媒体から読み出されたプログラムにより制御されるコンピュータは、請求項1の情報中継装置として機能する。

【0049】次に、請求項9に対応する発明は、請求項2に対応する発明をコンピュータに実現させるプログラムを記録した記録媒体である。

【0050】この記録媒体から読み出されたプログラムにより制御されるコンピュータは、請求項2の情報中継装置として機能する。

【0051】次に、請求項10に対応する発明は、請求項6に対応する発明をコンピュータに実現させるプログラムを記録した記録媒体である。

【0052】この記録媒体から読み出されたプログラムにより制御されるコンピュータは、請求項6の情報中継装置として機能する。

【0053】

【発明の実施の形態】以下、本発明の実施の形態について説明する。

(発明の第1の実施の形態) 図1は本発明の第1の実施形態に係る情報中継装置を適用したプロキシサーバを含むネットワークシステムの構成例を示すブロック図である。

【0054】このネットワークシステムは、公衆回線121を介して多数のサーバノード114やクライアントノード101が接続されてなるインターネットである。また、クライアントノード101はプロキシサーバノード103を介して公衆回線121に接続され、サーバノード114に対するコンテンツ要求をプロキシサーバ103に行うようになっている。このプロキシサーバノード103に本実施形態の情報中継装置122が組み込まれている。

【0055】本実施形態では、各ノード101、103、114が接続されている場合について説明するが、ネットワークが介在することは発明の必須要件ではなく、例えばクライアントノード101と情報中継装置が同一のコンピュータ内のソフトウェアプロセスとして存在することも可能である。

【0056】また、図1では各ノード101、103、114が1つしか記載されていないが、クライアントノード101及びサーバノード114は同一の構成を持つ複数の装置が存在していてもよい。さらに、各ノード101、103、114は、通信機能を有する計算機システムからなり、そのCPUやメモリ等のハードウェア資源、並びに制御プログラム等のソフトウェア資源から図1に示す種々の機能実現手段を構成する。また、図1においては機能実現に必要な各種のデータが示されているが、これらは図示しない各計算機内のメモリやハードディスク等の記憶手段(図示せず)に記憶されるものである。

【0057】以下に、上記各要素の構成について説明する。

【0058】クライアントノード101には、WWWブラウザ等の情報表示機能を有するコンテンツ表示部102が設けられる。コンテンツ表示部102は、クライアントからサーバに向けてのコンテンツ取得リクエスト117を出力するとともに、サーバから送られてきたコンテンツを表示する。

【0059】次に、サーバノード114は、WWWサーバ等の情報発信機能を有するコンテンツ送信部115及びコンテンツデータ116を備えている。

【0060】コンテンツ送信部115は、クライアントもしくはプロキシサーバ103からのリクエスト117もしくは118を受けて、対応するコンテンツを送信する。

【0061】コンテンツデータ116は、クライアントもしくは情報中継装置からのリクエスト117もしくは118を受けて、コンテンツ送信部115から送信されるコンテンツの集合であり、図示しない記憶手段に格納

されている。また、個別のコンテンツを識別する ID (WWWの場合はURL) とコンテンツ内容実体との組で1つのコンテンツが表現される。以下の説明では、IDがURLである場合について説明する。

【0062】図2はURLと送信元のIDの一例を示す図である。

【0063】同図に示す情報は、クライアントノード101が取得したいコンテンツを指定するリクエストデータ117や118に含まれる識別情報の形態例である。

【0064】次にプロキシサーバノード103は、クライアントノード101の代理としてインターネットとの通信等を行うサーバであって、本実施形態では、この中に情報中継装置122が設けられている。

【0065】情報中継装置122は、入力情報統計データ105、レプリケーション戦略知識107、レプリカコンテンツURL109、URL別レプリケーション間隔108及びレプリカコンテンツデータ110の各データを保持する。また、これらの各データ及びリクエストデータ117に基づき、所定の処理を実行する入力解析部104、レプリケーション戦略決定部107、レプリカ利用判定部112、リクエスト／コンテンツ転送部113及びレプリカ更新部111を備えている。

【0066】ここでまず、入力解析部104は、クライアントから送られたリクエスト117を解析し、入力情報統計データ105を作成する。また、リクエストの全情報をレプリカ利用判定部112に送る。

【0067】入力情報統計データ105は、図示しない記憶手段に過去のリクエストの解析情報等が保存されるものである。例えば図3(a)の項目に付いて、図4に示す形態でリクエストの統計情報を保持する。

【0068】図3は入力情報統計データ及びレプリケーション戦略決定部で使用するデータの内容例を示す図である。

【0069】図4は入力情報統計データの保存形態の一例を示す図である。

【0070】なお、図3(a)に示す各項目のうち、URL(1)、最新リクエスト発行時刻(2)、総リクエスト回数(3)、リクエスト発生頻度(4)については、入力解析部104が受信したリクエスト117に基づき、解析あるいは計算されて更新されるものである。また、レプリカ使用回数(5)については、後述するレプリカ利用判定部112によってレプリカが使用されたときにインクリメントされる。

【0071】レプリケーション戦略決定部106は、レプリケーション戦略知識107に基づいて入力情報統計データ105を利用し、どのコンテンツをレプリケーションの対象にし、どの程度の間隔でレプリカを最新状態に更新するか等を決定し、その結果をURL別レプリケーション間隔108に出力する。ここで、レプリケーション戦略知識107は、レプリケーション戦略決定部1

06で利用する知識データであり、例えば図5に示す形態で知識を保持する。

【0072】図5はレプリケーション戦略知識の内容例を示す図である。

【0073】同図に示すレプリケーション戦略知識107の各例に対応して、レプリケーション戦略決定部106がどのようなレプリケーション間隔決定を行うかを説明すると以下の通りである。

【0074】レプリケーション戦略決定部106は、各URLに対応するレプリカが図5に示す処理#1~#4のうち、何れかの処理条件を満たすか否かを調べ、条件を満たす場合には、その処理に規定されたレプリケーション間隔をURL別レプリケーション間隔108に出力する。また、レプリケーション登録削除の場合には、レプリカコンテンツURL109から対応するURLを削除する。

【0075】また、レプリケーション戦略決定部106は、レプリケーション戦略決定処理に先立ち、図3(b)に示すレプリカ使用率(6)を計算し、当該戦略決定処理の間、これを保持する。

【0076】さて、図5における処理#1は、最新リクエスト発行時刻(2)が現在時刻から3時間より前であって、かつ、レプリカ使用率(6)が0.5より大きい場合には、そのURLについてのレプリケーション間隔を5時間とするものである。この条件で比較的長時間である5時間と間隔設定するのは、リクエスト頻度の割には情報中継装置122が保持するレプリカは高確率でそのまま使用されることから、サーバノード114における対応コンテンツの変更頻度が相対的に低いと予想できるからである。別のいい方をすれば、この場合はコンテンツの変更頻度に比べてリクエスト頻度が相対的に高いということになる。

【0077】図5における処理#2は、最新リクエスト発行時刻(2)が現在時刻から3時間より前であって、かつ、レプリカ使用率(6)が0.5以下の場合には、そのURLについてのレプリケーション間隔を2時間とするものである。この条件で比較的短い時間である2時間と間隔設定するのは、リクエスト頻度の割には情報中継装置122が保持するレプリカの利用率は低く、サーバノード114における対応コンテンツの変更頻度が相対的に高いと予想できるからである。別のいい方をすれば、この場合はコンテンツの変更頻度に比べてリクエスト頻度が相対的に低いということになる。

【0078】図5における処理#3は、単位時間あたりのリクエスト発生頻度(4)が5より大きく、かつ、レプリカ使用率(6)が0.8より小さい場合には、そのURLについてのレプリケーション間隔を頻度計測単位時間の1/3とするものである。これは、レプリケーション間隔を単なる時間で設定するのではなくリクエスト発生頻度やレプリカ使用率といった概念に対応させて規

定するものである。また、この処理#3の場合は、リクエスト要求が高くかつサーバ側におけるコンテンツ変更頻度が高い場合に対応している。

【0079】図5における処理#4は、現在時刻から3日以上リクエストがない場合には、レプリケーションを終了させるものである。この場合は、レプリケーション登録削除の場合であり、レプリケーション戦略決定部106はレプリカコンテンツURL109、URL別レプリケーション間隔108及びレプリカコンテンツデータ110からそのURLについての関連情報を削除する。

【0080】なお、図5に示すレプリケーション戦略知識107は、種々考えられる戦略知識の例示であって、その他の戦略知識を適用させることもできる。例えば処理#4にAND条件で”固定でないコンテンツであるか否か”という条件を加え、これらの両条件が成立したときのみ”レプリケーション終了”とさせてもよい。この場合には、3日リクエストされずかつクライアントが固定的にリクエストするコンテンツでない場合にのみ、レプリケーションが登録削除されることになる。

【0081】URL別レプリケーション間隔108は、レプリケーション戦略決定部106で決定された、URL毎のレプリケーション間隔のデータが保持されたものである。例えば図6に示す形態でURL毎のレプリケーション間隔の値を保持する。

【0082】図6はURL別レプリケーション間隔の内容例を示す図である。

【0083】レプリカコンテンツURL109は、レプリケーション戦略決定部106で決定された、レプリケーション対象となるコンテンツのURLリストが保持されているものである。例えば図7に示すような形態でレプリケーション対象コンテンツのURLを保持する。

【0084】図7はレプリカコンテンツURLの一例を示す図である。

【0085】レプリカコンテンツデータ110は、レプリカコンテンツURL109にリストアップされているURLに対応するコンテンツのレプリカ実体が保持されたものである。例えば図8に示す形態でレプリカコンテンツ実体を保持する。

【0086】図8はレプリカコンテンツデータの一例を示す図である。

【0087】レプリカ更新部111は、URL別レプリケーション間隔108に記述されている間隔で各URLについてのレプリカコンテンツデータ110を更新するように、リクエスト/コンテンツ転送部113を介してサーバノードにコンテンツを要求する。

【0088】レプリカ利用判定部112は、リクエストデータ117に含まれる時刻と、レプリカコンテンツデータ110中のリクエストと対応するコンテンツ実体の有効期限もしくは更新時刻とを比較し、クライアントノード101に対してレプリカを返すかサーバノード11

4から新たに取得してデータを返すかを判定する。また、レプリカをクライアントノード101に返すように判断した場合には、対応するURLについて、入力情報統計データ105におけるレプリカ使用回数をインクリメントする。

【0089】リクエスト/コンテンツ転送部113は、レプリカ利用判定部112の判定結果に基づいて、サーバノード114にコンテンツ取得リクエスト118を送出し、サーバノード114から返されるコンテンツデータをクライアントノード101に返す。レプリカ利用判定部112にてレプリカを返すように判断されたときには、クライアントノード101にレプリカを返す。

【0090】また、リクエスト/コンテンツ転送部113は、サーバノード114から取得したコンテンツデータをレプリカコンテンツデータ110及びレプリカコンテンツURL109に格納する。したがって、リクエスト/コンテンツ転送部113は、レプリカ更新のみならず、新規にリクエストされたコンテンツの格納及びURL保持も行う。

【0091】なお、請求項におけるレプリケーション手段は、例えばレプリカ更新部111、URL別レプリケーション間隔108、レプリカコンテンツデータ110、レプリカコンテンツURL109及びリクエスト/コンテンツ転送部113等が対応し、複製返却手段は、レプリカ利用判定部112及びリクエスト/コンテンツ転送部113等が対応する。

【0092】次に、以上のように構成された本実施形態における情報中継装置の動作について、図9、図10及び図11を用いて説明する。

【0093】図9は本実施形態におけるリクエスト中継処理を説明する流れ図である。

【0094】図10は本実施形態におけるレプリケーション戦略決定処理を説明する流れ図である。

【0095】図11は本実施形態におけるレプリカ更新処理を説明する流れ図である。

【0096】これら各処理のうち、リクエスト中継処理及びレプリケーション戦略決定処理と、レプリカ更新処理とは独立に進む非同期な動作となる。

【0097】まずリクエスト中継処理について説明する。

【0098】このリクエスト中継処理は、図9に示すように、クライアントノード101からリクエストデータ117が発行された場合に主に情報中継装置122にて行われる処理である。

【0099】まず、コンテンツ表示部102に対して利用者がURLを指定することにより、コンテンツ表示要求がクライアントノード101から出力される(S1)。

【0100】利用者から出された上記表示要求は、リクエストデータ117の形態でコンテンツ表示部102か

ら情報中継装置122の入力解析部104へ送信される(S2)。

【0101】リクエストデータ117の内容は入力解析部104により解析され、リクエストを受け取った時刻の情報を元に、入力情報統計データ105の内容(1)～(4)が更新される(S3)。入力情報統計データ更新後は(S3)、後述するレプリケーション戦略決定処理へ進む。

【0102】一方、ステップS2の処理を受けて、入力解析部104を介して受信した要求内容に基づき、レプリカコンテンツデータ110に格納されている当該URLのレプリカコンテンツの有効期限がレプリカ利用判定部112によって算出される(S4)。

【0103】ここで、有効期限の算出方法は種々の方法が利用可能である。例えばサーバノード114からコンテンツが送信される際、有効期限時刻が添付されるか、最終変更時刻が添付されるため、その値から算出する方法が採用される。

【0104】次に、リクエストに対応するレプリカコンテンツが有効期限内であった場合(S5)、対応するレプリカコンテンツがレプリカコンテンツデータ113より読み出される。この読出コンテンツは、リクエスト/コンテンツ転送部113からクライアントノード101に返される(S6)。

【0105】さらに、レプリカ利用判定部112によって入力情報統計データ106におけるレプリカ使用回数(5)が"1"増やされ(S7)、レプリケーション戦略決定処理となる。

【0106】また、リクエストに対応するレプリカコンテンツが有効期限を過ぎていた場合、又はレプリケーションされていないコンテンツに対するリクエストである場合(S5)、リクエストデータ117に含まれるサーバノード名に基づいて、リクエスト/コンテンツ転送部113からサーバノード114にリクエストデータ118が転送される。サーバノード114からはリクエストデータ118で指定されたコンテンツが情報中継装置103に返される(S8)。

【0107】サーバノード114から返されたコンテンツは、情報中継装置122のリクエスト/コンテンツ転送部113からクライアントノード101に転送されるとともに(S9)、リクエスト/コンテンツ転送部113の処理により、レプリカコンテンツデータ110の内容が同コンテンツで更新される(S10)。なお、レプリケーションされていないコンテンツの場合には、新たなレプリケーション対象としてデータ(コンテンツ、URL)保持され、以降レプリケーションされることになる。

【0108】なお、レプリカコンテンツデータ110の内容が更新された後は(S10)、レプリケーション戦略決定へと処理が進む。

【0109】次に、レプリケーション戦略決定処理の流れについて説明する。レプリケーション戦略決定処理はリクエスト中継処理の図9；S3若しくはS7、S10の処理を受けて開始する。

【0110】まず、図9のステップS3若しくはS7、S10の後に、入力情報統計データ105及びレプリケーション戦略知識107をもとに、レプリケーション戦略決定部106によってURL別レプリケーション間隔108が生成される(t1)。

【0111】具体的には、例えば図4のような入力情報統計データと図5のようなレプリケーション戦略知識の場合には以下のようにレプリケーション間隔が決定される。

【0112】まず、リクエストデータ117で指定されるURLの入力情報統計データが取り出される。例えばリクエストデータでhttp://www.foo.htmlが指定された場合、図4の1行目の情報が取り出される。

【0113】次に、レプリケーション戦略決定部106によってレプリカ使用率(6)が計算されるとともに、取り出した入力情報統計データにマッチする条件を持つ知識がレプリケーション戦略知識107から検索される。例えば現在時刻が1998/09/07/10:00であった場合、図4の1行目のデータについて検索すると、図5の1行目の知識がマッチする。

【0114】次に、見つかった知識に記されているレプリケーション間隔を当該URLのレプリケーション間隔とし、URL別レプリケーション間隔108がレプリケーション戦略決定部106によって更新される。

【0115】以上が新たなレプリケーション間隔を設定する場合である。なお、他の処理に該当する場合については上記したとおりである。

【0116】これに対して、見つかった知識に記されているレプリケーション間隔がレプリケーションの終了を示している場合には、レプリケーション戦略決定部106によって、当該URL及び対応するレプリカコンテンツデータがURL別レプリケーション間隔108、レプリカコンテンツURL109、レプリカコンテンツデータ110の各データから削除される(t2)。

【0117】また、レプリケーションを終了させる他の方法として、レプリカコンテンツデータ110が定められた大きさよりも大きくなった場合には、コンテンツの優先度情報(図示せず)を参照し、必要性の低いコンテンツをレプリケーション対象から除外するようにしてもよい。この優先度情報としては、例えば最新リクエスト発行時刻が用いられる。また、使用者が別途優先度を設定するようにしてもよい。さらに、この場合には、優先度の低いレプリカを削除する、プロキシサーバノード103のレプリカ削除機能を利用し、レプリカコンテンツ等を削除するようにしても良い。

【0118】なお、本実施形態ではルールベースによるレプリケーション間隔の決定方法を説明したが、入力情報統計データとレプリケーション戦略知識の形態はこれに限るものではなく、入力情報統計データの変量解析によるレプリケーション間隔値の算出等も考えられる。

【0119】次に、レプリカ更新処理の流れについて説明する。レプリカ更新処理はリクエスト中継処理及びレプリケーション戦略決定処理とは非同期に動作するものである。

【0120】まず、レプリカコンテンツURL109に格納されている最終レプリカ作成時刻と、URL別レプリケーション間隔108に格納されているレプリケーション間隔とがレプリカ更新部111により比較される。この結果、現在時刻が最終レプリカ作成時刻からレプリケーション間隔の時間以上経過しているURLがあれば、そのコンテンツをサーバノード114からダウンロードするように、レプリカ更新部111からリクエスト／コンテンツ転送部113に要求される。

【0121】リクエスト／コンテンツ転送部113によってコンテンツがダウンロードされ、レプリカコンテンツデータ110がダウンロードデータで置き換えられ、レプリカが更新される(u1)。

【0122】レプリカ更新後、レプリカ更新部111は一定時間停止し、その後、ステップu1の処理が繰り返される(u2)。

【0123】上述したように、本発明の実施の形態に係る情報中継装置は、レプリカ更新部111を設け、リクエスト中継処理と非同期にレプリケーション処理を実行して、予めレプリカコンテンツを更新しておくようにしたので、情報中継装置内のレプリカコンテンツデータが最新のものに保たれている確率を高くすることができ、クライアントリクエストに対するヒット率を向上させることができる。

【0124】また、レプリケーション戦略決定部106を設け、リクエストの頻度に応じてレプリカ対象コンテンツを増減させるようにしているので、限られた容量のレプリカ記憶領域に対し、クライアントリクエスト頻度の高いコンテンツがレプリケーションされるようになり、クライアントからの不特定なコンテンツ取得要求に対するヒット率を向上させることができる。

【0125】また、今までのPushプラットフォーム(Castanet等)では、あらかじめ設定したコンテンツだけがPushの対象となり、クライアント側の要求の変化に柔軟に対応できないが、本発明では、レプリカコンテンツデータ110に入っていないコンテンツは、以降自動的にPush対象になるため、柔軟なキャッシング処理(レプリケーション)を実現させることができる。

【0126】さらに、今までのPushプラットフォームでは、キャッシュエリアが足りなくならないように、

必要なくなったコンテンツを全て手作業でPush対象から外す必要があった。これに対して、本発明では、コンテンツ毎に設定される優先度に基づいて自動的にPush対象から外すようにしたので、キャッシュエリア(レプリカコンテンツデータ110)が溢れることがない。

(発明の第2の実施の形態) 第1の実施形態では、レプリカ利用判定部112にてレプリカが有効期限を過ぎていと判定された場合には、リクエスト117を無条件にサーバノード114へリクエスト118として転送していた。

【0127】これに対し、本実施形態では、レプリカコンテンツがサーバノード上のオリジナルコンテンツと同一であるかどうかを確認することにより、サーバノード114からの無駄なコンテンツのダウンロードを無くすようにしたものである。

【0128】本実施形態の情報中継装置は、リクエスト／コンテンツ転送部113の処理が修正される他、第1の実施形態と同様に構成されている。

【0129】リクエスト／コンテンツ転送部113は、第1の実施形態と同様に構成される他、サーバノード114からコンテンツをダウンロードする際、まず、コンテンツ属性情報のみを取得し、当該コンテンツ属性情報に基づきレプリカコンテンツが最新でない場合にのみコンテンツの全体を取得するようになっている。

【0130】このためのリクエストデータ118の内容として、図12、図13に示すように2種類が用意されている。

【0131】図12は本発明の第2の実施形態にて使用されるコンテンツ属性情報だけを取得するリクエスト例を示す図である。

【0132】同図のリクエストでは、リクエストのタイプが「属性情報のみ」であって、そのURLが「http://www.aaa.bbb/foo.html」であることを示している。

【0133】ここで、コンテンツの属性情報としては、少なくとも「コンテンツの最終変更時刻(新規コンテンツの場合は作成時刻)」が含まれている。

【0134】図13はコンテンツ全体を取得するリクエスト例を示す図である。

【0135】同図のリクエストでは、リクエストのタイプが「コンテンツ全体」であって、そのURLが「http://www.aaa.bbb/foo.html」であることを示している。

【0136】また、リクエスト／コンテンツ転送部113は、入力情報統計データ105にオリジナルコンテンツの変更頻度情報を追加する。この情報は例えば以下の式によって求められる。

【0137】変更頻度情報＝コンテンツの全体取得リクエスト回数÷コンテンツの属性取得リクエスト回数

レプリケーション戦略知識 107 は、オリジナルコンテンツの変更頻度を条件に含むことができる。例えばオリジナルコンテンツの変更頻度が所定値範囲の条件であれば、レプリケーション間隔をある値に設定する等の処理である。

【0138】次に、以上のように構成された本実施形態における情報中継装置の動作について、図 14 を用いて説明する。

【0139】図 14 は本実施形態におけるリクエスト中継処理を説明する流れ図である。

【0140】本実施形態の動作においては第 1 の実施形態と比較してリクエスト中継処理が変更される。

【0141】まず、レプリカコンテンツが有効期限内である場合には、第 1 の実施形態と同様な処理が実行される (v1 ~ v7)。すなわち図 9 に示すステップ S1 ~ S7 と、ステップ v1 ~ v7 とは同様であるため、その説明は省略する。

【0142】次に、レプリカが有効期限を過ぎていると判定された場合 (v5)、リクエスト/コンテンツ転送部 118 によりサーバノード 114 にコンテンツ属性取得リクエスト 118 が送信され、当該コンテンツの属性情報が取得される (v8)。

【0143】取得した属性情報に含まれる最終変更時刻と、対応するレプリカの最終レプリカ作成時刻とがリクエスト/コンテンツ転送部 113 にて比較され、オリジナルコンテンツがレプリカより新しいかが判定される (v9)。

【0144】オリジナルコンテンツがレプリカより新しいと判定された場合 (v9)、サーバからコンテンツ全体が取得され (v10)、レプリカコンテンツデータ 110 が更新され (v11)、さらに、当該 URL に関する入力情報統計データが変更される。

【0145】また、コンテンツ全体が取得された後 (v10)、当該コンテンツがクライアントノードに返される (v12)。

【0146】一方、ステップ v9 において、最終レプリカ作成時刻がオリジナルコンテンツの最終変更時刻より遅い場合、レプリカは最新のオリジナルコンテンツと同一であると判定され、レプリカコンテンツがクライアントノード 101 へ返される (v13)。その後、レプリカ使用回数がインクリメントされる (v14)。

【0147】上述したように、本発明の実施の形態に係る情報中継装置は、レプリカコンテンツがサーバノード上のオリジナルコンテンツと同一であるかどうかを確認するようにしたので、サーバノード 114 からの無駄なコンテンツのダウンロードを無くすことができ、有効期限を過ぎても変化が無いコンテンツに対するレスポンス時間の短縮が期待できる。

【0148】また、オリジナルコンテンツの更新頻度に基づきレプリケーション間隔を調整することが可能であ

るので、有効期限が明示的に設定されていないコンテンツのレプリカ更新間隔を確率的な意味で適切に設定することができ、リクエスト中継処理中にコンテンツ全体を取得しなければならない状況が起きにくくなる。

【0149】なお、本実施形態では属性情報の事前要求をクライアントノード 101 からリクエスト 117 があつた場合についてのみ説明したが、この属性情報の事前要求は、リクエスト中継処理のみならず、レプリカ更新処理の場合に適用しても良い。この場合、レプリカ更新部 111 がレプリカ更新すべきと判断したときに、属性情報の事前要求が行われ、コンテンツが変更しているときのみダウンロードされることになる。

(発明の第 3 の実施の形態) 第 1、第 2 の実施形態は、いわゆるインテリジェントプルと呼ばれる技術を応用したものであるが、本実施形態ではプッシュ技術を利用する。このプッシュ技術利用と関連し、リクエスト/コンテンツ転送部とコンテンツ送信部の間の通信路にマルチキャスト技術が利用されており、これによりサーバノードから複数の情報中継装置に対し能動的にデータを送信することができるようになっている。

【0150】図 15 は本発明の第 3 の実施形態に係る情報中継装置を適用したプロキシサーバを含むネットワークシステムの構成例を示すブロック図である。なお、同図に示す情報中継装置 203 はプロキシサーバに含まれるものであるが、両者の関係は第 1 の実施形態の場合と同様であるので、ここではプロキシサーバの表示を省略している。

【0151】本実施形態に用いられるネットワーク自体や、クライアントノード 201、プロキシサーバノード及びサーバノード 211 間の関係、さらにはこれら各部の基本的な機能は、第 1 の実施形態の各ノード 101、103 及び 114 の場合と同様であるのでここでは詳細説明を省略する。

【0152】まず、クライアントノード 201 はクライアントノード 101 と同一の構成である。また、クライアントノードから送られるリクエストデータ 208 もリクエストデータ 117 と同一である。

【0153】情報中継装置 203 は、レプリケーション対象 URL 206、レプリカコンテンツデータ 207、レプリケーション戦略知識 221 及び入力情報統計データ 222 の各データを図示しない記憶手段に保持する。また、これらの各データ及びリクエストデータ 208 に基づき、所定の処理を実行するレプリカ利用判定部 204、リクエスト/コンテンツ転送部 205 及び登録削除判断部 223 を備えている。

【0154】ここでまず、レプリカ利用判定部 204 は、リクエストデータ 208 に含まれる URL のコンテンツがレプリケーション対象 URL 206 に存在すればレプリカコンテンツデータ 207 から当該コンテンツをクライアントノードに返し、存在しなければレプリケー

10

20

30

40

50

ション対象URL 206に当該URLを追加し、サーバノード211に新規レプリケーション対象コンテンツとして登録するリクエスト209を送る。

【0155】また、レプリカ利用判定部204は、リクエストデータ208に基づき、そのURL別に、最新リクエスト発行時刻を入力情報統計データ222として保持させる。

【0156】リクエスト／コンテンツ転送部205は、サーバノード211から送られて来たコンテンツデータ210のうち、レプリケーション対象URL 206に登録されているURLのコンテンツのみをレプリカコンテンツデータ207に記録する。

【0157】レプリケーション対象URL 206は、情報中継装置203がレプリケーション対象とするコンテンツのURLを記録するデータエリアである。例えば図16の形態のデータ構造を持つ。

【0158】図16は本実施形態におけるレプリケーション対象URLのデータ構造例を示す図である。

【0159】レプリカコンテンツデータ207は、レプリカコンテンツを記録するデータエリアである。例えば図17の形態のデータ構造を持つ。

【0160】図17は本実施形態におけるレプリカコンテンツデータのデータ構造例を示す図である。

【0161】レプリケーション戦略知識221には、レプリカコンテンツデータ207内のレプリカコンテンツを削除する条件を示すデータである。この条件は、「最新のリクエスト発生が現在時刻の3日前」や「最新のリクエスト発生が現在時刻の3日前 AND 当該URLが固定コンテンツでない」等である。

【0162】登録削除判断部223は、入力情報統計データ222の最新リクエスト発行時刻及びレプリケーション戦略知識221の登録削除条件から登録削除すべきURLを判断する。同判断部223は、あるレプリカコンテンツを登録削除すべきと判断したときには、レプリケーション対象URL 206及びレプリカコンテンツデータ207から対応するデータを削除するとともに、サーバノード211に登録削除データを送信する。

【0163】一方、サーバノード211は、コンテンツデータ214、レプリケーション対象URL 216及びレプリケーション配送先装置名217の各データを図示しない記憶手段に保持するとともに、これらのデータに関連して一定の処理を行うコンテンツ送信部212、更新検出部213及びレプリケーション対象URL登録部215を備えている。

【0164】レプリケーション対象URL登録部215は、リクエストデータ209に含まれるURLをレプリケーション対象URL 216に登録するとともに、リクエストデータ209の送信元の情報中継装置名をレプリケーション配送先装置名217に登録する。

【0165】また、レプリケーション対象URL登録部

215は、情報中継装置203の登録削除判断部223から削除要求データ224を受けると、当該データ224に含まれるURLをレプリケーション対象URL 216から削除し、対応するデータをレプリケーション配送先装置名217から削除する。

【0166】レプリケーション対象URL 216は、サーバノード211から情報中継装置203にレプリケーションのために送信するコンテンツのURLリストである。複数の情報中継装置203からリクエストデータ209を受けている場合、各情報中継装置を区別せず記録する。例えば情報中継装置のレプリケーション対象URL 206と同様な図16の形態を持つ。

【0167】コンテンツデータ214は、サーバノード211がオリジナルデータを持つコンテンツデータの集合である。

【0168】更新検出部213は、コンテンツデータ214の変化を検知し、それがレプリケーション対象URL 216に含まれていれば、URLとコンテンツデータ実体をコンテンツ送信部212へ送る。

【0169】更新検出部213におけるコンテンツデータの変化を検知する方法としては、例えば以下のような何種類かの方法が考えられる。

【0170】A. 一定時間毎に全てのコンテンツの最終更新時刻を確認し、前回の確認時刻よりも後に更新が起きたコンテンツを見つける。

【0171】B. コンテンツを変更する時に使用するツールから、変更後のコンテンツをコンテンツデータ214に記録する際に更新検出部213に知らせる。

【0172】C. OSやハードウェアで、記憶装置上でデータの変化があったことを検出する機能が備わっている場合、データ変化を検出した時点で更新検出部213に知らせる。

【0173】コンテンツ送信部212は、更新検出部213から送られたURLとコンテンツの組を、レプリケーション配送先装置名217に登録されている全ての情報中継装置203に向けて送信する。

【0174】次に、以上のように構成された本実施形態における情報中継装置の動作について、図18、図19、図20及び図21を用いて説明する。

【0175】図18は本実施形態におけるリクエスト中継処理を説明する流れ図である。

【0176】図19は本実施形態におけるレプリカ対象登録処理を説明する流れ図である。

【0177】図20は本実施形態におけるレプリカ配信処理を説明する流れ図である。

【0178】図21は本実施形態における登録削除処理を説明する流れ図である。

【0179】ここで、図18及び図19の処理は一連の流れとして処理が進むが、図20及び図21は独立した処理の流れとなる。

【0180】まず、図18、図19の動作の流れを説明する。

【0181】ステップw1は図9のステップS1と同一の処理になる。

【0182】次に、クライアントノード201からのリクエストデータ208は、レプリカ利用判定部204にて受信され(w2)、リクエストに含まれるURLがレプリケーション対象URL206に登録済みかが調べられる(w3)。

【0183】既に登録済みであった場合には(w3)、レプリカ利用判定部204によってレプリカコンテンツデータ207のレプリカがクライアントノード201に返され(w4)、終了する。

【0184】一方、登録済みでなかった場合(w3)、レプリケーション対象URL206に、リクエストに含まれるURLが追加される(w5)。さらに、サーバノード211に対し、レプリケーション対象URL登録及びコンテンツ取得のリクエスト209が送信される(w6)。

【0185】サーバノード211のコンテンツ送信部212によって、リクエスト209に含まれるURLのコンテンツが情報中継装置203に返され、さらに情報中継装置203によりクライアントノード201に転送される。(w7)このリクエスト中継処理に連動してレプリカ対象登録処理が図19に示すようにして実行される。

【0186】まず、リクエスト209に含まれるURLはレプリケーション対象URL登録部215に引き渡され、このURLがサーバノード211のレプリケーション対象URLに追加される(x1)。さらに、リクエスト209の送信元装置の識別名がレプリケーション対象URL登録部215によりレプリケーション配信先装置名217に追加される(x2)次にレプリカ配信処理について説明する。

【0187】まず、コンテンツデータ214内のコンテンツの内容変更もしくは新規作成が更新検出部213で検知される(y1)。

【0188】変化を検知したコンテンツがレプリケーション対象URL216に含まれるコンテンツか否か更新検出部213により判断される(y2)。

【0189】ここで、レプリケーション対象であった場合には、コンテンツ送信部212を通じてレプリケーション配信先装置名217に登録されている装置にコンテンツ210が送信される(y3)。この送信の際には、マルチキャスト技術が利用され、これによりサーバノードから複数の情報中継装置203に対し能動的にデータ送信が行われる。図15に示すのはこの複数の情報中継装置203のうちの一つである。

【0190】情報中継装置203においては、サーバノード211から送られて来たコンテンツ210のURL

が、レプリケーション対象URL206に含まれるかどうかリクエスト/コンテンツ転送部205によって判定される(y4)。レプリケーション対象URLに含まれる場合には、レプリカコンテンツデータ207はコンテンツ210によって更新される(y5)。

【0191】このように、レプリカ対象登録処理によって、レプリカ対象がプッシュ技術を利用したサーバノード211に登録され、さらに、その登録URLについてレプリカ配信処理によって情報中継装置203に配布される。

【0192】ここで、サーバノード211に登録されるレプリケーション対象URLが増えすぎないようにするために、情報中継装置203によって登録削除処理が行われる。

【0193】図21に示す登録削除処理では、まず、登録削除判断部223によって登録削除判断が開始され、そのために、入力情報統計データ222の最新リクエスト発行時刻及びレプリケーション戦略知識221の登録削除条件が読み出される(z1)。

【0194】つづいて、最新リクエスト発行時刻に基づき、各URLについて、登録削除条件を満たしているかが判断される(z2)。

【0195】その結果、登録削除条件を満たすURLがあれば、そのURLに関連したレプリケーション対象URL206及びレプリカコンテンツデータ207が削除され(z3)、サーバノード211に対して削除要求データ224が出力される(z4)。

【0196】サーバノード211においては、削除要求データ224を受け取ったレプリケーション対象URL登録部215により、同データ224に含まれるURLについてのレプリケーション対象URL216及びレプリケーション配信先装置名217についての情報が削除される(z5)。

【0197】上述したように、本発明の実施の形態に係る情報中継装置は、サーバノード211におけるオリジナルコンテンツデータの変更と連動してレプリケーション処理が起きようになっているので、情報中継装置中のレプリカコンテンツデータは常に最新状態に保たれ、クライアントノードからのリクエストに迅速に対応することができる。

【0198】すなわち今までのProxyサーバは一度取得したコンテンツはそのまま保持するだけなので、オリジナルが更新されたら、次のブラウザからのアクセス時にもう一度サーバから転送し直さなければならなかった。しかし本発明では、常にProxyの持つコンテンツが最新であるため、ブラウザから見た平均レスポンス時間の向上を見込むことができる。

【0199】なお、本実施形態では、サーバノード211上でコンテンツデータ214の変更が起きた時点で変更されたコンテンツ自体を情報中継装置203に送信す

るようにしたが、本発明はこのような形態に限られるものではない。例えば第2の実施形態と組み合わせ、コンテンツデータ214の変更時にはコンテンツ自体でなく、その属性情報のみをリクエスト／コンテンツ転送部205に送信する。そして、同転送部205にてレプリカコンテンツが最新のものであるか否かを判断した上で、レプリカが最新でない場合にのみ、リクエスト／コンテンツ転送部205から同コンテンツのリクエスト118を発行し、レプリカを更新するようにしてもよい。

(発明の第4の実施の形態) 第1、第2及び第3の実施形態では、コンテンツのURL及び実体のみにについてのレプリケーションについて説明しているが、本実施形態では、コンテンツ毎に設定されているアクセスコントロール情報をレプリカコンテンツと共に保存する場合を説明する。

【0200】本実施形態の情報中継装置は、アクセスコントロール判断部が設けられ、レプリカコンテンツデータ110、207にアクセスコントロール情報が保持されるようになっている他、第1～第3の実施形態と同様に構成されている。

【0201】ここで、アクセスコントロール情報とは、「Aというユーザはこのコンテンツを見ることができる」といった内容の情報であり、これによりクライアントノードの利用者毎にアクセスできるコンテンツを制限することができる。

【0202】アクセスコントロール判断部は、このアクセスコントロール情報に基づいて、コンテンツアクセス制限を実行する。

【0203】このように構成された本実施形態の情報中継装置の動作としては、まず、クライアントからのリクエストデータにユーザ識別名が加わる。

【0204】情報中継装置内のレプリカコンテンツデータに登録されたアクセスコントロール情報と、クライアントユーザ識別名とがアクセスコントロール判断部により比較される。その結果、アクセスが許されたユーザにのみレプリカコンテンツが返される。

【0205】上述したように、本発明の実施の形態に係る情報中継装置は、コンテンツと共にアクセスコントロール情報を付加して配布するようにしたので、アクセスコントロールが必要なコンテンツに対しても、クライアントから見た時のレスポンスタイムを短縮することができる。

【0206】なお、同様な場合、第1～第3の実施形態では、ユーザ毎のアクセスコントロール情報を得るためにはサーバノードに問い合わせる必要がある。

(発明の第5の実施の形態) 上記第1～第4の実施形態は、情報中継装置をインターネットのプロキシサーバに適用する例で説明したが、本実施形態ではファイルサーバシステムに適用する場合を説明する。

【0207】図22は本発明の第5の実施形態に係る情

報中継装置を適用するファイルサーバシステムの全体構成例を示す図である。

【0208】このシステムは、本社301に大規模ファイルサーバ302が設置され、これが第1～第4の実施形態のサーバノード114、211に対応する。この大規模ファイルサーバ302は、比較的低速なLANを介して本社301内の各部門に設けられた小規模ファイルサーバ303、304に接続され、また公衆回線を介して各支社の小規模ファイルサーバ305、306に接続されている。

【0209】これらの小規模ファイルサーバ303～306に情報中継装置122、203が設けられる。

【0210】各小規模ファイルサーバ303～306は、比較的小規模の容量のファイルサーバであり、高速LANでその部門若しくは支社内のクライアントにファイル情報を提供できるようになっている。また、自己が有しない情報を要求された場合には、同情報を大規模ファイルサーバ302に請求する。この情報が第1～第4実施形態のコンテンツに相当する。また、ファイル検索情報がURLに相当する。

【0211】また、大規模ファイルサーバ302は、全ての情報を格納するものであり、小規模ファイルサーバ303～306からの要求に応じてデータを送信するようになっている。

【0212】このように構成されたファイルサーバシステムでは、クライアント～小規模ファイル～大規模ファイルサーバ間が、第1～第4の実施形態におけるクライアントノード～プロキシサーバノード～サーバノード間と同様な関係となる。

【0213】したがって、その部門や支社においてある程度以上の頻度で使用する情報は、容量の少ない小規模データベースに格納され、使用しない情報は小規模データベースから排除されていくことになる。

【0214】上述したように、本発明の実施の形態に係る情報中継装置は、第1～第4の実施形態と同様な構成をファイルサーバシステムに設けたので、元のファイルサーバに、セクション毎に設けられた小ファイルサーバが問合せに行くような場合でも、第1～第4の実施形態と同様な効果が得られる。

【0215】また、本発明の適用範囲は、インターネットやファイルサーバシステムに限られるものでなく、実施形態に示した状況と類似する状況が生じる計算機・通信システムに対しては種々適用することが可能である。例えばFTP等にも適用可能である。

【0216】さらに、実施形態に説明した装置は、記憶媒体に格納したプログラムをコンピュータに読み込ませることで実現させることができる。

【0217】ここで本発明における記憶媒体としては、磁気ディスク、フロッピーディスク、ハードディスク、光ディスク(CD-ROM、CD-R、DVD等)、光

10

20

30

40

50

磁気ディスク（MO等）、半導体メモリ等、プログラムを記憶でき、かつコンピュータが読み取り可能な記憶媒体であれば、その記憶形式は何れの形態であってもよい。

【0218】また、記憶媒体からコンピュータにインストールされたプログラムの指示に基づきコンピュータ上で稼働しているOS（オペレーティングシステム）や、データベース管理ソフト、ネットワークソフト等のMW（ミドルウェア）等が本実施形態を実現するための各処理の一部を実行してもよい。

【0219】さらに、本発明における記憶媒体は、コンピュータと独立した媒体に限らず、LANやインターネット等により伝送されたプログラムをダウンロードして記憶又は一時記憶した記憶媒体も含まれる。

【0220】また、記憶媒体は1つに限らず、複数の媒体から本実施形態における処理が実行される場合も本発明における記憶媒体に含まれ、媒体構成は何らの構成であってもよい。

【0221】なお、本発明におけるコンピュータは、記憶媒体に記憶されたプログラムに基づき、本実施形態における各処理を実行するものであって、パソコン等の1つからなる装置、複数の装置がネットワーク接続されたシステム等の何れの構成であってもよい。

【0222】また、本発明におけるコンピュータとは、パソコンに限らず、情報処理機器に含まれる演算処理装置、マイコン等も含み、プログラムによって本発明の機能を実現することが可能な機器、装置を総称している。

【0223】

【発明の効果】以上詳記したように本発明によれば、オリジナルコンテンツの更新頻度が高く、さらにクライアントから不特定多数のサーバへの情報提供要求が出力される場合において、その平均レスポンスタイムを向上させることができる情報中継装置及び記憶媒体を提供することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態に係る情報中継装置を適用したプロキシサーバを含むネットワークシステムの構成例を示すブロック図。

【図2】URLと送信元のIDの一例を示す図。

【図3】入力情報統計データ及びレプリケーション戦略決定部で使用されるデータの内容例を示す図。

【図4】入力情報統計データの保存形態の一例を示す図。

【図5】レプリケーション戦略知識の内容例を示す図。

【図6】URL別レプリケーション間隔の内容例を示す図。

【図7】レプリカコンテンツURLの一例を示す図。

【図8】レプリカコンテンツデータの一例を示す図。

【図9】同実施形態におけるリクエスト中継処理を説明する流れ図。

【図10】同実施形態におけるレプリケーション戦略決定処理を説明する流れ図。

【図11】同実施形態におけるレプリカ更新処理を説明する流れ図。

【図12】本発明の第2の実施形態にて使用されるコンテンツ属性情報だけを取得するリクエスト例を示す図。

【図13】コンテンツ全体を取得するリクエスト例を示す図。

【図14】同実施形態におけるリクエスト中継処理を説明する流れ図。

【図15】本発明の第3の実施形態に係る情報中継装置を適用したプロキシサーバを含むネットワークシステムの構成例を示すブロック図。

【図16】同実施形態におけるレプリケーション対象URLのデータ構造例を示す図。

【図17】同実施形態におけるレプリカコンテンツデータのデータ構造例を示す図。

【図18】同実施形態におけるリクエスト中継処理を説明する流れ図。

【図19】同実施形態におけるレプリカ対象登録処理を説明する流れ図。

【図20】同実施形態におけるレプリカ配信処理を説明する流れ図。

【図21】同実施形態における登録削除処理を説明する流れ図。

【図22】本発明の第5の実施形態に係る情報中継装置を適用するファイルサーバシステムの全体構成例を示す図。

【符号の説明】

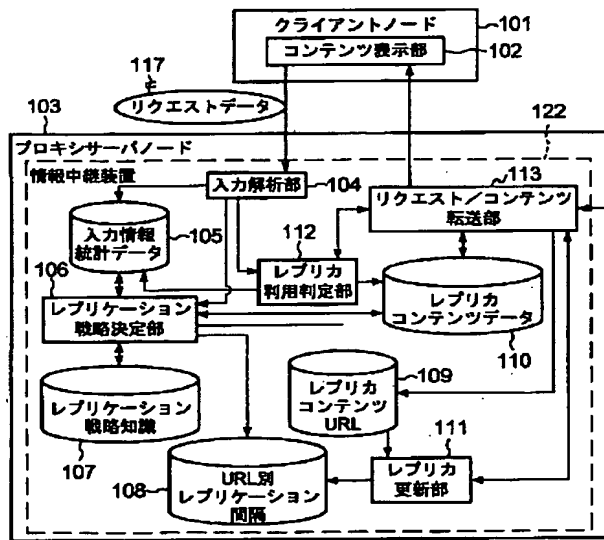
- 101…クライアントノード
- 103…プロキシサーバノード
- 104…入力解析部
- 105…入力情報統計データ
- 106…レプリケーション戦略決定部
- 107…レプリケーション戦略知識
- 108…URL別レプリケーション間隔
- 109…レプリカコンテンツURL
- 110…レプリカコンテンツデータ
- 111…レプリカ更新部
- 112…レプリカ利用判定部
- 113…リクエスト/コンテンツ転送部
- 114…サーバノード
- 115…コンテンツ送信部
- 116…コンテンツデータ
- 117, 118…コンテンツ取得リクエスト
- 121…公衆回線
- 122…情報中継装置
- 201…クライアントノード
- 202…コンテンツ表示部
- 203…情報中継装置

204…レプリカ利用判定部
 205…リクエスト/コンテンツ転送部
 206…レプリケーション対象URL
 207…レプリカコンテンツデータ
 208…リクエストデータ
 209…リクエストデータ
 210…コンテンツデータ
 211…サーバノード
 212…コンテンツ送信部
 213…更新検出部
 214…コンテンツデータ

215…レプリケーション対象URL登録部
 216…レプリケーション対象URL
 217…レプリケーション配送先装置名
 221…レプリケーション戦略知識
 222…入力情報統計データ
 223…登録削除判断部
 301…本社
 302…大規模ファイルサーバ
 303, 304, 305, 306…小規模ファイルサーバ

10 バ

【図1】



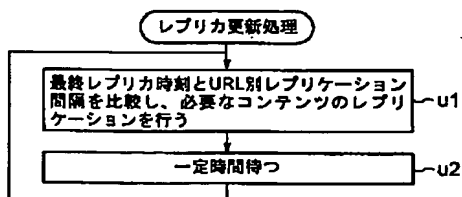
【図2】

URL	http://www.toshiba.co.jp/foo/bar.html
送信元	client1.toshiba.co.jp

【図4】

(1)	(2)	(3)	(4)	(5)
http://www/foo.html	1998/09/07/09:12	35	3	22
http://www2/bar.gif	1998/09/07/08:00	16	7	12

【図11】



【図16】

URL
http://www/foo.html
http://www2/bar.gif

【図3】

(a) 入力情報統計データ

項目番号	項目名	項目の意味
(1)	URL	コンテンツを特定するURL
(2)	最新リクエスト発行時刻	最も最近に当該URLについて発行されたリクエストの発行時刻
(3)	総リクエスト回数	システム管理者が定めた時点から現在までの、当該URLに関する総リクエスト数
(4)	リクエスト発生頻度	単位時間当たりのリクエスト発生回数
(5)	レプリカ使用回数	レプリカが返された数

(b) レプリケーション戦略決定部での計算データ

(6)	レプリカ使用率	リクエスト数に対し、レプリカが返された数の割合 (5)/(3)
-----	---------	---------------------------------

【図12】

項目	内容
リクエストタイプ	属性情報のみ
URL	http://www.aaa.bbb/foo.html

【図 5】

処理	条件	レプリケーション間隔
#1	{最新リクエスト発行時刻 (2) > 現在時刻の3時間} AND {レプリカ使用率 (6) > 0.5}	5 時間
#2	{最新リクエスト発行時刻 (2) > 現在時刻の3時間} AND {レプリカ使用率 (6) ≤ 0.5}	2 時間
#3	{リクエスト発生頻度 (4) > 5} AND {レプリカ使用率 (6) < 0.8}	頻度計測単位時間/3
#4	最新リクエスト発行時刻 (2) > 現在時刻の3日前	レプリケーション終了

【図 6】

URL	レプリケーション間隔
http://www/foo.html	5 時間
http://www2/bar.gif	1 時間

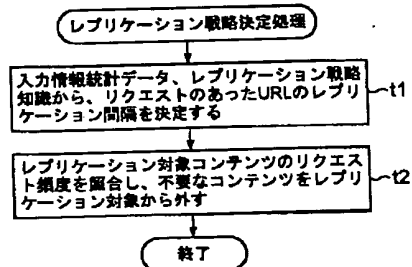
【図 7】

URL	最終レプリカ作成時刻
http://www/foo.html	1998/09/07/09:00
http://www2/bar.gif	1998/09/07/09:00

【図 8】

URL	レプリカ実体	有効期限
http://www/foo.html	foo.htmlを構成するデータ	1998/09/08/08:10
http://www2/bar.gif	bar.gifを構成するデータ	1998/09/08/09:35

【図 10】



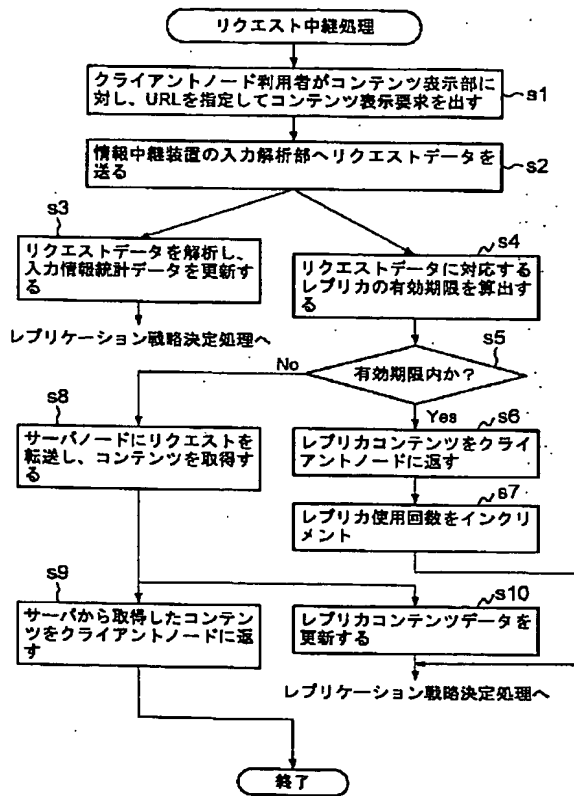
【図 13】

項目	内 容
リクエストタイプ	全体
URL	http://www.aaa.bbb/foo.html

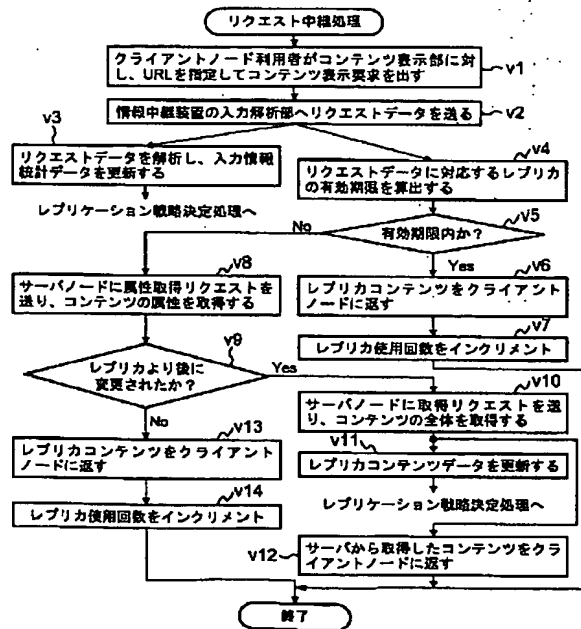
【図 17】

URL	レプリカ実体
http://www/foo.html	foo.htmlを構成するデータ
http://www2/bar.gif	bar.gifを構成するデータ

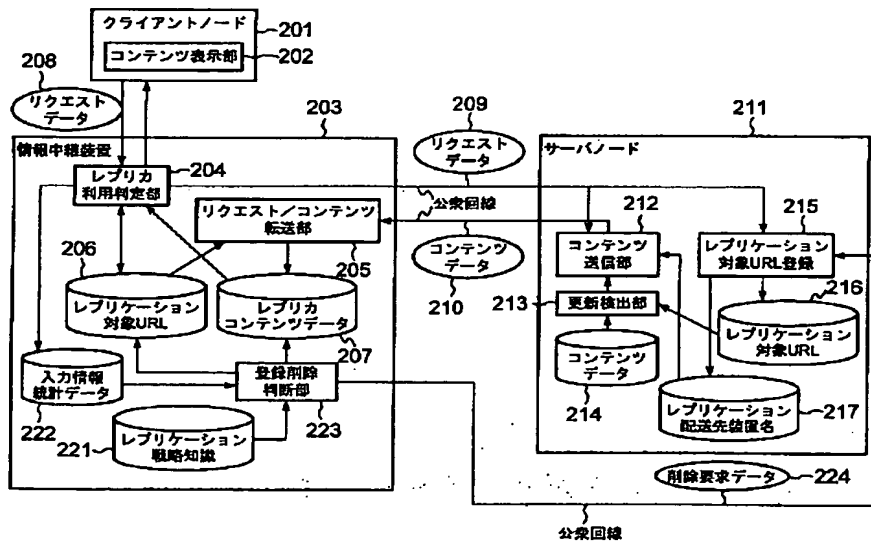
【図 9】



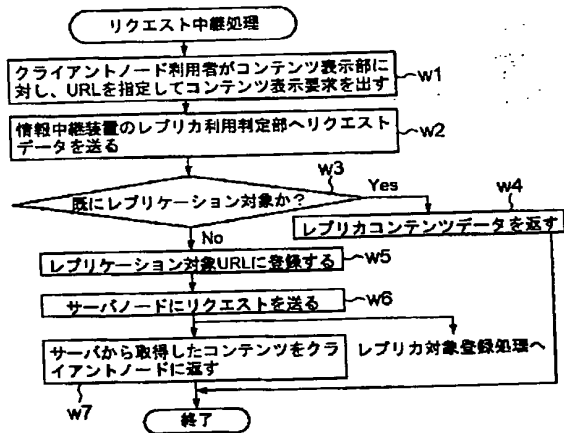
【図 14】



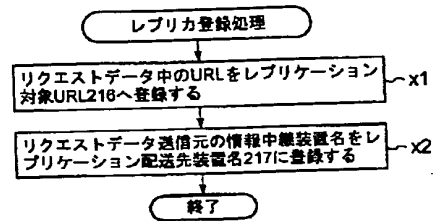
【図 15】



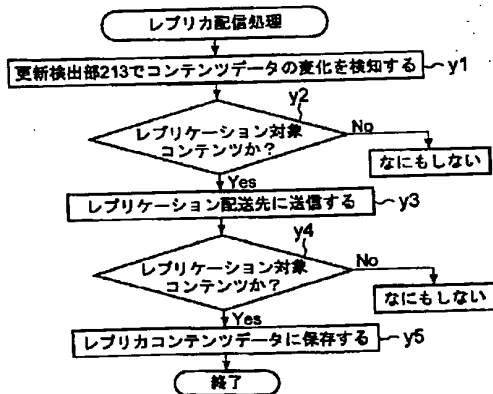
【図18】



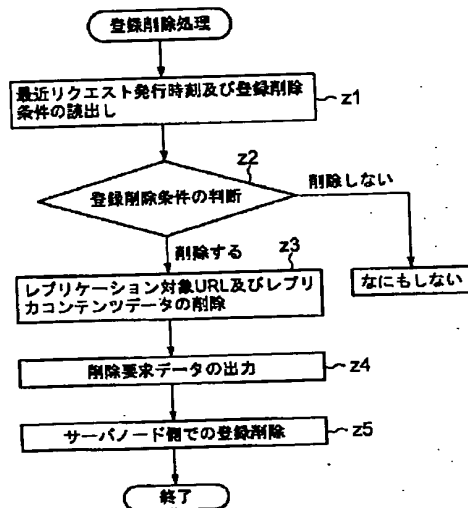
【図19】



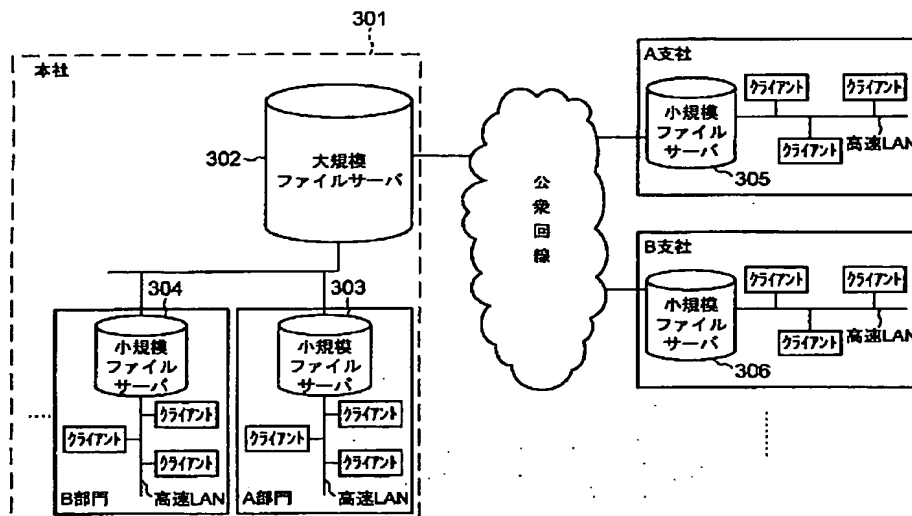
【図20】



【図21】



【図 22】



フロントページの続き

Fターム(参考) 5B089 GA19 GA21 HA01 JA22 JB15
 KA05 KB06 KD04
 5K030 GA03 HA04 HC01 HD01 KA07
 LD17 LD20
 9A001 JJ04 JJ25 JJ27

THIS PAGE BLANK (USPTO)